2017 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)
September 24–28, 2017, Vancouver, BC, Canada

Automatic Extrinsic Calibration Of Depth Sensors With Ambiguous Environments And Restricted Motion

Jarrett Holtz[1] and Joydeep Biswas[2]

***Abstract*— Autonomous mobile robots that use multiple depth sensors to perceive their environments, rely on extrinsic calibration to combine the individual views from each sensor into a single coherent view of the surroundings. Such extrinsic calibration is tedious to perform manually, and requires that specific scenes to calibrate. Current state of the art automatic approaches do not consider the content of scenes used for calibration, and thus are not robust to partially informative scenes in long-term deployments. In this paper, we present Delta-Calibration, an automated extrinsic calibration technique that takes into account the information in a scene for calibration. Delta-Calibration relies on constrained sensor motion to minimize the effects of desynchronization, and ego-motion estimation from each depth camera to detect significant changes in pose, which we term Delta-Transforms. We derive a solution to the extrinsic calibration using such Delta-Transforms taking into account uncertain axes of motion in the environment, and further infer necessary and sufficient conditions on the Delta-Transforms such that Delta-Calibration results in a unique, non-singular, and numerically stable extrinsic calibration. We present quantitative and qualitative results demonstrating the effectiveness of Delta-Calibration at computing extrinsic calibration over different arrangements of depth sensors.**

I. INTRODUCTION

Inexpensive depth cameras such as the Microsoft Kinect sensor have proven to be invaluable for mobile robot autonomy, and have been used for various tasks such as mapping, autonomous indoor mobile robot localization, human detection, and object identification. However, due to the limited field of view of such depth sensors, mobile robots need to be equipped with multiple depth sensors. In order to assemble a single coherent view of the environment from a system of multiple depth sensors, the robot must rely on the pose transforms for each sensor to transform their observations into a single consistent reference frame. The problem of estimating the set of pose transforms for the sensors is called extrinsic calibration.

There have been many proposed approaches to extrinsic calibration of sensors, but existing state of the art approaches perform poorly when calibrating in scenes with few features, which we call partially informative scenes. In order for calibration to be possible in diverse environments and long-term deployment scenarios it needs to be robust to environments with varying amounts of visual information. In this paper, we present an approach to automatic unsupervised extrinsic calibration of depth sensors that places no restrictions on the arrangement of sensors, requires no engineered information, and can calibrate from multiple partially informative scenes. We call our approach Delta-Calibration since it relies on the perceived ego-motion, or Delta-Transforms of each depth sensor when rigidly connected to each other. Delta-Calibration computes the extrinsic calibration as the solution to a system of linear equations in the special Euclidean group in 3D, $SE(3)$. We derive the solution of the extrinsic calibration in analytic closed-form in terms of the observed Delta-Transforms in Section III we use the derivation of the solution to demonstrate requirementes of the ego-motion based solution to extrinsic calibration, and based on these requirements develop a solution for calibrating ground-based robots. Given the analytic solution we develop a solution to the extrinsic calibration with uncertain ego-motions measured from partially informative scenes in Section IV. Finally, we present quantitative results of Delta-Calibration, and comparison to other methods using several configurations of example depth sensor arrays, and depth sensor to odometry calibration which show that Delta-Calibration performs better than the state of the art in partially informative scenes and with restricted axes of robot motion in Section V.

II. RELATED WORK

Related work for Delta-Calibration includes calibration methods for both intrinsics and extrinsics of depth and color sensors. The most common calibration technique is the supervised checkerboard method [21], which uses a checkerboard as an engineered target for intrinsic calibration of standard cameras, but other supervised methods have been employed that use single bright spots [17], cuboids [7], or spheres [15]. Supervised methods compute the calibration between the two images using co-planar checkerboard points identified in the color image [20], by using RANSAC for estimation [12], by implementing unique calibration targets for the Kinect and the color camera individually [10], or by calibrating the two simultaneously to correct for distortion [4]. Supervised methods require engineered targets, and as such are infeasible for calibration of robots operating in long phases of autonomy.

Unsupervised methods remove the need for engineered targets, and thus require less human interaction. Some unsupervised approaches use properties of the environment, such as grid maps [6]. Others rely on the motion of the objects in the environment by tracking planar correspondence points [16], by utilizing the geometry of large planes which can be seen without overlap by multiple sensors [2], by tracking

[1]Jarrett Holtz and [2]Joydeep Biswas are with the College of Information and Computer Sciences, University of Massachusetts, Amherst, MA 01003, USA {jaholtz}{joydeepb}@cs.umass.edu.

This work is supported in part by AFRL and DARPA under agreement #FA8750-16-2-0042.

correspondence points on arbitrary moving objects [13], or via nonparameteric learning using a mutual information model [9]. These methods rely on knowledge about the environment or specific qualities of scenes.

Methods more similar to our approach use motion of the sensors for calibration. This includes using sensor motion to calibrate 2D and 3D LIDARs by optimizing the point cloud with respect to a measure of entropy [11]. Most similar to our method are approaches which use sensors in motion. Some of these perform intrinsic calibration for depth sensors [18], [19], or with stereo vision [5] using rigid body motions for calibration. Others compute extrinsics using shared features between the sensors, either with bundle adjustment [3], or by utilizing the output of SLAM [1]. Work utilizing sensor ego-motion in a similar formulation to ours exists that handles uncertainty in the sensor readings, but not uncertainty in the scene or restricted axis of motion, by solving an uncertainty weighted version of the Kabsch algorithm for the extrinsic transform [18], [19]. Others use the Unscented Kalman Filter to estimate the pose between odometry enabled sensors, or constrained dual quaternion optimization to solve the motion relationship between the sensors [14].

Delta-Calibration is an unsupervised method that calculates the extrinsic calibration based on observed motion from each sensor. We place no restriction on the field of view, and require no calibration targets. Further, Delta-Calibration is robust to partially informative scenes, and restricted axes of sensor motion. Removing these requirements makes Delta-Calibration applicable to real-world autonomous calibration scenarios in varied environments.

III. EXTRINSIC CALIBRATION USING DELTA-TRANSFORMS

Extrinsic calibration relates two sensors via the 3-Dimensional rigid-body transform that takes a point from one camera's frame of reference to its corresponding location in the second sensor's frame of reference. We denote the coordinate frame for sensor i at time t as C_i^t. For a given sensor we refer to the affine transform that takes points from sensor $i's$ frame of reference at timestep $t+\Delta t$ to timestep t as a *DeltaTransform*, written as D_i^t. Delta-Calibration seeks to take corresponding sets of Delta-Transforms from a rigidly connected pair of sensors, and from them compute the extrinsic calibration between the two sensors.

Given C_i^t and $C_i^{t+\Delta t}$ for $i = \{1, 2\}$, the Delta-Transforms D_1^t, and D_2^t, and the affine transform A, there are two orderings of transforms that can take points from the reference frame $C_2^{t+\Delta t}$ to the reference frame C_1^t:

$$D_1^t A p_2^{t+\Delta t} = A D_2^t p_2^{t+\Delta t} = p_1^t. \tag{1}$$

In general, the Delta-Transforms themselves are related as $D_1^t A = A D_2^t$. This is a special case of the equation form called the Sylvester Equation [8] where the solution is an element of the Special Euclidean group $SE(3)$. To solve this equation, we consider the homogeneous matrix representation of the transforms where each transform is composed of a rotation matrix and a translation as,

$$D_1^t = \begin{bmatrix} R_1 & T_1 \\ 0 & 1 \end{bmatrix}, \; D_2^t = \begin{bmatrix} R_2 & T_2 \\ 0 & 1 \end{bmatrix}, \; A = \begin{bmatrix} R & T \\ 0 & 1 \end{bmatrix} \tag{2}$$

Thus, the rotation and translation components are expressed separately as

$$R_1 R = R R_2, \quad R T_2 + T = R_1 T + T_1. \tag{3}$$

The rotation component of Eqn. 3 is again a special case of the Sylvester equation where the solution is an element of the special orthogonal group $SO(3)$. We now present how Eqn. 3 is solved to recover the rotation R, and using this computed result we recover the translation T.

A. Rotation Calculation

We first note that a unique solution to Eqn. 3 exists if, and only if rotations R_1 and R_2 are not co-axial. If R_1 and R_2 are co-axial, there exist an infinite family of solutions with a rotation axis parallel to R_1 and R_2, and with arbitrary magnitude of rotation. To calculate the extrinsic rotation R we first represent Eqn. 3 in the equivalent quaternion form $q_1 q = q q_2$, where each rotation R is represented by its equivalent quaternion form, such that rotation R corresponds to quaternion $q = w + xi + yj + zk$, R_1 to $q_1 = w_1 + x_1 i + y_1 j + z_1 k$, and R_2 to $q_2 = w_2 + x_2 i + y_2 j + z_2 k$. From the rules of quaternion multiplication, we then convert into the linear system $M_1 q_v = M_2 q_v$, where $q_v = [w \; x \; y \; z]^T$ is the vector form of the equivalent quaternion q, and matrices M_1 and M_2 are given by the rules of quaternion post-multiplication by q_2 and pre-multiplication by q_1 respectively,

$$M_1 = \begin{bmatrix} w_1 & -x_1 & -y_1 & -z_1 \\ x_1 & w_1 & -z_1 & y_1 \\ y_1 & z_1 & w_1 & -x_1 \\ z_1 & -y_1 & x_1 & w_1 \end{bmatrix}, \tag{4}$$

$$M_2 = \begin{bmatrix} w_2 & -x_2 & -y_2 & -z_2 \\ x_2 & w_2 & z_2 & -y_2 \\ y_2 & -z_2 & w_2 & x_2 \\ z_2 & y_2 & -x_2 & w_2 \end{bmatrix}. \tag{5}$$

Thus, the quaternion vector q_v corresponding to the rotation R must then lie in the null space of $(M_1 - M_2)$. By computing the eigenvalue decomposition of $M_1 - M_2$ we find q_v to be the unit-norm eigenvector corresponding to the lowest eigenvalue. The resulting q_v is then the extrinsic rotation R in quaternion form. Rotation can be found for multiple Delta-Transforms in quaternion vector form q_v for all M_1^t and M_2^t at time t as

$$q_v = \arg\min_{q_v} \sum_t [(M_1^t - M_2^t) q_v]^T [(M_1^t - M_2^t) q_v]. \tag{6}$$

The solution for q_v can then be found by computing the eigenvalue decomposition of the combined matrix $\mathbf{M}$ from all time-steps, where $\mathbf{M}$ is given by,

$$\mathbf{M} = \sum_t [(M_1^t - M_2^t)]^T [(M_1^t - M_2^t)]. \tag{7}$$

Given the value of R, we can then solve for T as described next.

B. Translation Calculation

Having calculated the rotation component R of the extrinsic calibration A, to calculate the translation component T, we first rearrange the equation for translation as $(I-R_1)T = T_1 - RT_2$. and solve for T.

$$T = (a^T a)^{-1} a^T b, \;\; a = (I - R_1), \;\; b = T_1 - RT_2. \quad (8)$$

Eqn. 8 has no solution for a single rotation matrix R_1, since for any proper rotation matrix R, $(I - R)v = 0$ for all v parallel to the axes of rotation of R. Thus $a^T a$ is not invertible for a single pair of observed delta transforms, so at least two rotations with non-parallel axis must be used in the extrinsic calculation. Therefore, a solution to T using multiple DeltaTransforms is computed from,

$$T = \left(\sum_t a_t^T a_t\right)^{-1} \left(\sum_t a_t^T b_t\right). \quad (9)$$

Here, the values of a_t, b_t are computed from the Delta-Transforms from timestep t by Eqn. 8. Using Eqn. 9 we can calculate the extrinsic translation T between the two sensors, and the combination of R and T gives us the full extrinsic calibration A. Thus, given at least two Delta-Transforms per sensor with non-parallel axes of rotation from rigidly connected sensors which are not coaxial we can compute the extrinsic calibration between the two sensors.

C. Calculating Delta-Transforms

In order for Delta-Calibration to yield quality calibration results accurate Delta-Transforms need to be calculated. These transforms can be calculated in a number of ways, but the set of Delta-Transforms should observe certain properties for robust calibration. These qualities are: 1) Delta-Transforms must be time-aligned across sensors, 2) they should be calculated over keyframes in order to yield transforms distinct from sensor variance, 3) and the set of transforms should cover all axes of motion for the sensor system.

D. Calibration with restricted motion

We have presented a solution to calibration given sets of multiple Delta-Transforms which contain rotations with non-parallel axes. However, when the robot has restricted degrees of freedom, we cannot guarantee multiple axes of rotation. With this restriction Eqn. 7 and Eqn. 9 can yield only partially correct solutions. This is because Delta-Transforms with only z-axis rotation contain no information about the z-axis of either R or T. We propose a solution for the z-axis of R given a pure translation from the robot, and for the z-axis of translation given a view of the ground plane and R, in the scenario where the desired calibration is between the sensor and the robot base.

Specifically, when $R_1 = I$ the translation equation simplifies to $RT_2 = T_1$. This has no solution for a single pair of translations T_2, T_1, nor for multiple pairs of translations which are co-axial. In order to solve for R we can use this in combination with Eqn. 3 to form a fully constrainted problem Delta-Transform pairs P_r and P_t such that P_r contains any combination of rotations and translations and P_t contains only pure translations, then a least squares problem can be formed which minimizes the residuals r_1 and r_2 described in Eqn. 10.

$$r_1 = \sum_i^{i \in P_r} (R_1^i - RR_2^i)^2, \quad r_2 = \sum_i^{i \in P_t} (RT_2^i - T_1^i)^2 \quad (10)$$

Given R we can calculate the x and y components of T using Eqn. 9, and we can calculate the z component given a point cloud which contains a view of the ground plane, and the partial extrinsic transform A. By transforming the point cloud with A we treat the distance from the sensor to the ground plane as the z-axis component of T. With these additional steps we can calibrate using ego-motion even in the special case of a robot with limited degrees of freedom.

IV. Calibration from Partially Informative Scenes

We have derived an analytical solution to extrinsic calibration from sensor ego-motion that holds when the ego-motions can be accurately determined. However, for depth sensors it is possible to have partially informative scenes, such as a view of a single wall, in which no features are present that can disambiguate change along specific axes. It is possible to determine the amount of information in a scene using distribution of normal vectors in a scene in the scene. Given the set of normal vectors for a scene $\mathbb{N} = \{n_i\}$ from $i = 0$ to $i = s$, we form the scatter matrix S and calculate its condition number $k(S)$ as:

$$S = \sum_i^s n_i n_i^T, \quad k(S) = \frac{\sigma_{max}(S)}{\sigma_{min}(S)} \quad (11)$$

Where $\sigma_{max}(S)$ is the largest singular value of S and $\sigma_{min}(S)$ is the smallest singular value of S. Then, if S is well conditioned such that $k(S)$ is close to 1 the scene is fully informative, otherwise some axis of motion will be ambiguous.

Given the normals we can determine the axis of translational certainty c_t or translational uncertainty u_t, and the axis of rotational unceraintty u_t. Given $\mathbb{N}$ we set u_r and c_t equal to the first normal n_0 because a single normal in a scene only gives information about translation along the normal, and gives no information about rotation about an axis parallel to the normal. Then for each additional normal n_i in $\mathbb{N}$ we check if n_i is parallel to u_r, if it is not there is no axis of rotational uncertainty. Second, if we have no single axis u_t, and c_t is not parallel to n_i, then u_t is the axis of translation orthogonal to c_t and n_i, and if we then have a single u_t then there is no single axis c_t. Finally, if we had some single u_t and u_t is not orthogonal to n_i then we have no axis of translational uncertainty u_t.

A. Calibration with Ambiguous Delta-Transforms

When some part of the Delta-Transforms cannot be measured from a partially informative scene, the relationship described in Eqn. 1 will not hold. We propose an extension given known axis of translational uncertainty u_{it}, and known axis of rotational uncertainty u_{ir}, for all sensors $i = \{1, 2\}$.

B. Rotation Calculation

To calculate the extrinsic rotation R we start with the quaternion representation of our rotation calculation, where q represents the quaternion form of a rotation. In order to account for the uncertainty we need to find the component of q_1 corresponding to axis u_{2r}, q_1^{u2}, and the component of q_2 corresponding to axis of uncertainty u_{1r}, q_2^{u2}. To do this we calculate the portion of the angle-axis forms $\vec{q_1}, \vec{q_2}$ of the respective quaternions, corresponding to the axes of uncertainty using the extrinsic rotation R, and then we treat the resultant angle axes rotations $q_1^{\vec{u2}}, q_2^{\vec{u1}}$ as new quaterions q_1^{u2} and q_2^{u1}. We can calculate these angle-axis rotations as follows:

$$
\begin{aligned}
q_1^{\vec{u2}} &= (R^{-1}\vec{q_1} \cdot u_{r2})u_{r2}, \\
q_2^{\vec{u1}} &= (R\vec{q_2} \cdot u_{r1})u_{r1}.
\end{aligned} \tag{12}
$$

Where $\cdot$ corresponds to the vector dot product. Given this, A solution to rotation considering uncertainty can be written as:

$$q_1 q q_1^{u2^{-1}} = q_2^{u1^{-1}} q q_2 \tag{13}$$

$$R^* = \arg\min_q \sum_i q_1^i q q_1^{u2^i{}^{-1}} - q_2^{u1^i{}^{-1}} q q_2^i \tag{14}$$

The solution for q can then be calculated using least-squares optimization, and using the rotation matrix form of q, R, we can then solve for T.

C. Translation Calculation

To calculate T we rewrite Eqn. 8 and set it equal to 0.

$$E_t = T - R_1 T - T_1 - RT_2 \tag{15}$$

Eqn. 15 is the total error with a given T and some uncertainty. To account for uncertainty we can remove the portion of the error vector E_t along the axis of uncertainty u_{1t} and u_{2t} as:

$$E = (E_t) - (E_t \cdot u_{1t}^i)u_{1t}^i - (E_t \cdot Ru_{2t}^i)Ru_{2t}^i \tag{16}$$

In order to fully account for possible sources of uncertainty in T we must still acount for the uncertainty in R_1, u_{1r}. This is because rotational uncertainty implies that a portion of translation which was resultant of rotation could be unobserved. Given this we can calculate the residual error given some T as $r = (E \cdot u_1^i)^2$.

Which we use to calculate a least-squares solution to the problem given R_1^i, t_1^i, and t_2^i for all timesteps i as.

$$T^* = \arg\min_T \sum_i r^i \tag{17}$$

Where T^* is the extrinsic translation.

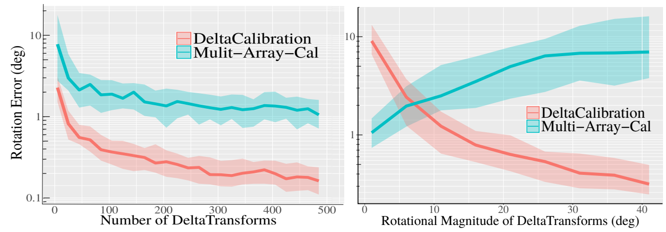


Fig. 1: Effects of varying experimental parameters of Delta-Transforms on rotational error of calibration with no ambiguity or restricted motion.

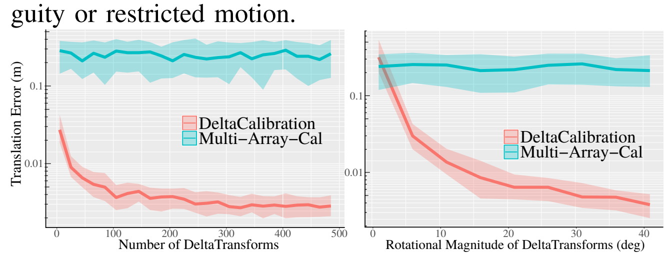


Fig. 2: Effects of varying experimental parameters of Delta-Transforms on translation error of calibration with no ambiguity or restricted motion.

D. Calibration with restricted motion

We need a special case to handle a robot which can only rotate about the z axis. In order to handle uncertainty we must adjust the equation for translation to handle u_{1t} and u_{2t}. The restructured equation is as follows:

$$RT_2 - R(T_2 \cdot R^{-1}u_{1t})R^{-1}u_{1t}) = T_1 - (T_1 \cdot Ru_{2t})Ru_{2t} \tag{18}$$

Eqn. 19 can be solved using multiple Delta-Transforms for the estimated transform R^*.

$$r^i = RT_2^i - R(T_2^i \cdot R^{-1}u_{1t}^i)R^{-1}u_{1t}^i) - T_1^i - (T_1^i \cdot Ru_{2t}^i)Ru_{2t}^i$$

$$R^* = \arg\min_R \sum_i r^i \tag{19}$$

Given this calculation for the z-axis of R given restricted motion and partial information we can calculate the full extrinsics using Eqn. 14 and Eqn. 17.

V. Experimental Results

We performed three sets of experiments to evaluate Delta-Calibration with respect to other calibration techniques: 1) simulation experiments to evaluate the effect of noise in, the magnitude of, and number of Delta-Transforms used for Delta-Calibration, and the ego-motion based calibration technique described in [18] which we will call Multi-Array-Cal; 2) real-world calibration experiments s; and 3) real-world calibration experiments with a Kobuki Turtlebot and partially informative scenes.

A. Simulation Experiments

To evaluate the effect of sensor noise on the accuracy of Delta-Calibration, we generate synthetic sequences of Delta-Transforms $R_1^t, T_1^t, R_2^t, T_2^t$, for a randomly chosen extrinsic calibration R, T. Here, R_i^t, T_i^t are the rotation, and translation components of the Delta-Transforms for sensor

i at time-step $t, t \in [1, N]$. We generate synthetic sequence of Delta-Transforms from randomly generated ground-truth extrinsic calibration R, T, and evaluate calibration by comparing resultant extrinsic transforms $\hat{R}, \hat{T}$ to R, T.

We perform three types of simulation experiments over three different conditions and compare the results of Delta-Calibration and Multi-Array-Calib. The experiments: 1) use Delta-Transforms with unambigous and unrestricted motion, 2) with only ambigous motion, and 3) with motion restricted to only planar x,y plane translations and z-axis rotations. The entire procedure is repeated 100 times for each condition. The three conditions we test are 1) the effect of the magnitude of noise in rotation ϵ_R, ϵ_T, 2) the magnitude of the Delta-Transforms E_R, E_T, and 3) the number of Delta-Transforms N used for Delta-Calibration. While varying each parameter we hold the others at nominal values $N = 10, \epsilon_R = 10°, \epsilon_T = 2\text{cm}, E_R = 1°, E_T = 1\text{cm}$.

Fig. 1 and Fig. 2 show the plots of error in extrinsics as calculated by Delta-Calibration. The first graphs show that both methods improve with additional number's of Delta-Transforms, but the closed form solution of Delta-Calibration outperforms regardless of the number of Delta-Transforms. The results of Multi-Array-Cal in these experiments are comparable to the results presented in [18], [19], but it should be noted that these papers used the output of Multi-Array-Cal as initial guesses for further optimizations that require data from RGB sensors, which we do not have for our experiments.

In the second set of graphs in Fig. 1 and Fig. 2 we see that Multi-Array-Cal outperforms Delta-Calibration for very small Delta-Transforms. This is related to the handling of sensor variance as part of the calibration in Multi-Array-Cal as described in [18], [19]. However, Delta-Calibration performs as well or better in all cases with lower information to noise ratios, and can also calibrate in scenarios which Multi-Array-Cal cannot.

The results of our experiments using Delta-Transforms with ambiguous and restricted axes of motion show similar results to Fig. 1 and Fig. 2 for Delta-Calibration and are ommitted for brevity. These cases are not handled by Multi-Array-Cal, and so the error and variance are both very high. In both cases Delta-Calibration performs with median errors of less than one degree of rotational error, while the Multi-Array-Cal method shows errors of as much as 100 degrees. This flexibility makes Delta-Calibration less brittle and more applicable to real robot scenarios.

B. Real-World Experiments

For real-world experiments we compare the results of different calibration techniques over multiple two sensor datasets. These datasets are:

1) Right-Angle, a view with a roughly 90° rotation between the two sensors with no overlap,
2) Wide-Horizontal, a non-overlapping vertical wide-angle view with *only* depth images, and no covariance estimate.
3) Overlap, an overlapping horizontal wide-angle view,
4) Opposite, a view with slightly less than 180° rotation between the two sensors which *never* view the same scene, and
5) Three-Sensors, a view with three sensors, two with overlap, with no shared scenes.

We calibrate these setups using different calibration techniques:

1) DeltaCal ICP: Iterative closest point based calculation of Delta-Transforms. Requires only depth images.
2) DeltaCal SLAM Raw: Delta-Calibration using Delta-Transforms calculated using SLAM keyframes and poses. Requires depth and color images.
3) DeltaCal SLAM Selected: Delta-Calibration using Delta-Transforms calculated using SLAM poses and Delta-Calibration keyframes. Requires depth and color images.
4) SLAM-Optimization: using an implementation of the SLAM based optimization method described in [1]. Requires depth images, color images, and views of the same features.
5) Multi-Array-Cal: using an implementation of the ego-motion based method which takes into account the variance of the sensor data described in [18], [19].

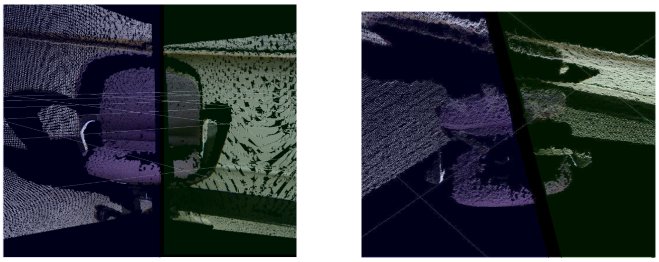
To evaluate the quality of final calibration we define an error metric to use given the absence of ground truth.

Consistency Error: Consistency error is a measure of the inconsistency in utilizing resultant extrinsic calibration to construct a 3d model. We measure this error using datasets that contain known planar features. We compute the angular consistency error (ACE) as the absolute value of the difference in rotation between the corresponding plane normals and the known relationship between the planes. We then compute the translational consistency error (TCE) as the absolute value of the difference between the point to plane distance and the known distance. We use three non-parallel planes per dataset in order to take into account all degrees of freedom, and we average the ACE and TCE from all planes. Table I shows the results of this error metric for our datasets, where an entry of — represents that a method could not be run on the given dataset.

For our datasets only ego-motion based methods could calibrate all datasets. From this data it can be seen that performance between Delta-Calibration using either selected keyframes from SLAM or DeltaCalculation yields consistency error with centimeters of translational magnitude and an average of .06 radians of ACE in the worst case. Delta-Transforms calculated via SLAM selected keyframes performed particularly poorly. In our tests the SLAM-Optimzation has variable performance based on the number of shared features viewed by the sensors, yielding more accurate results the closer the maps formed by SLAM are to identical. These results show that Delta-Calibration yields accurate calibration results for varied sensor setups, and importantly, that only Delta-Calibration can perform calibration in all of our test scenarios with reasonable results. As an example calibration Fig. 3 demonstrates an alignment

TABLE I: Error Metric Results for multiple sensor configurations and calibration methods. ACE in radians, and TCE in meters. Boxed columns represent results from our method.

SETUP	DeltaCal ICP		Delta-Cal SLAM Selected		SLAM-Optimization		Delta-Cal SLAM Raw		Multi-Array-Cal	
	TCE	ACE	TCE	ACE	TCE	ACE	TCE	ACE	TCE	ACE
Overlap	.034	.06	.018	.06	.353	1.09	.0325	.06	0.3	0.2
Right-Angle	.075	.04	.013	.055	.12	1.14	.043	.3	0.2	0.047
Opposite	.056	.06	.031	.05	1.67	1.16	—	—	0.64	0.07
Three-Sensors	.077	.054	.051	.06	.2215	1.05	—	—	0.425	0.05
Wide-Horizontal	.03	.026	—	—	—	—	—	—	—	—



(a) Chair Front View (b) Chair Top View

Fig. 3: Point clouds from wide-angle sensor configuration with no overlap, aligned using Delta-Calibration calculated extrinsics. The green and blue tinting of the point clouds differentiates data from the two sensors.

of a 3D scene calculated using Delta-Calibration.

C. Turtlebot Experiments

We performed two sets of experiments on the Kobuki Turtlebot platform. One with fully informative scenes, and one without. We calculated extrinsics with both Delta-Calibration and Multi-Array-Cal, we compared a set of fully informative Delta-Transforms recorded from the setup and recorded the average rotational and translation error in Table II. The results show that Delta-Calibration handles both cases without increase in error, and the partial scene case yields reasonable calibration results when the Multi-Array-Cal method cannot.

TABLE II: Error Metric Results for Turtlebot experiments. AEavg in radians, and TEavg in meters

SETUP	DeltaCal ICP		Multi-Array-Cal	
	TEavg	AEavg	TEavg	AEavg
Full-Scenes	.01	.04	.02	.16
Partial-Scenes	.01	.05	.10	.15

VI. Conclusion

In this paper we presented Delta-Calibration, a solution to extrinsic calibration of rigidly connected sensors using only motion observed from the sensors in the form of Delta-Transforms. Further our experiments have shown that Delta-Calibration calculates accurate extrinsic calibrations, even when presented with partially informative data or restricted motion which render other methods infeasible. In particular we have shown that Delta-Calibration is the only method in the state of the art which can perform extrinsic calibration of depth sensors, in partially informative environments with restricted robot motion. This makes Delta-Calibration well suited to the task of recalibrating robots in the field during phases of long-term autonomy.

References

[1] G. Carrera, A. Angeli, and A. J. Davison. SLAM-Based Automatic Extrinsic Calibration of a Multi-Camera Rig. In *ICRA*, pages 2652–2659, 2011.

[2] E. Fernndez-Moral, J. Gonzlez-Jimnez, P. Rives, and V. Arvalo. Extrinsic Calibration of a Set of Range Cameras in 5 Seconds Without Pattern. In *IROS*, pages 429–435, 2014.

[3] L. Heng, B. Li, and M. Pollefeys. CamOdoCal: Automatic Intrinsic and Extrinsic Calibration of a Rig with Multiple Generic Cameras and Odometry. In *IROS, 2013*, pages 1793–1800, 2013.

[4] D. Herrera C., J. Kannala, and J. Heikkil. Joint Depth and Color Camera Calibration with Distortion Correction. *Pattern Analysis and Machine Intelligence*, 34(10):2058–2064, 2012.

[5] R. Horaud, G. Csurka, and D. Demirdijian. Stereo calibration from Rigid Motions. *Pattern Analysis and Machine Intelligence*, 22(12):1446–1452, 2000.

[6] Y. Ji, A. Yamashita, and H. Asama. Automatic Calibration and Trajectory Reconstruction of Mobile Robot in Camera Sensor Network. In *Automation Science and Engineering (CASE)*, pages 206–211, 2015.

[7] B. Jin, H. Lei, and W. Geng. *ECCV*, chapter Accurate Intrinsic Calibration of Depth Camera with Cuboids, pages 788–803. 2014.

[8] B. Kågström and L. Westin. Generalized Schur Methods With Condition Estimators for Solving the Generalized Sylvester Equation. *Automatic Control*, 34(7):745–751, 1989.

[9] K. Kodagoda, A. Alempijevic, J. Underwood, S. Kumar, and G. Dissanayake. Sensor Registration and Calibration using Moving Targets. In *ICARCV*, pages 1–‘, 2006.

[10] W. Liu, Y. Fan, Z. Zhong, and T. Lei. A New Method for Calibrating Depth and Color Camera Pair Based on Kinect. In *ICALIP*, pages 212–217, 2012.

[11] W. Maddern, A. Harrison, and P. Newman. Lost in Translation (and Rotation): Rapid Extrinsic Calibration for 2D and 3D LIDARs. In *ICRA*, pages 3096–3102, 2012.

[12] I. Mikhelson, G. Lee, A. Sahakian, Y. Wu, and A. Katsaggelos. Automatic, Fast, Online Calibration Between Depth and Color Cameras . *Journal of Visual Communication and Image Representation*, 25(1):218 – 226, 2014.

[13] S. Miller, A. Teichman, and S. Thrun. Unsupervised Extrinsic Calibration of Depth Sensors in Dynamic Scenes. In *IROS*, pages 2695–2702, 2013.

[14] N. Roy, P. Newman, and S. Srinivasa. *Extrinsic Calibration from Per-Sensor Egomotion*, pages 504–. MIT Press, 2013.

[15] A. Staranowicz, G. R. Brown, F. Morbidi, and G. L. Mariottini. *PSIVT*, chapter Easy-to-Use and Accurate Calibration of RGB-D Cameras from Spheres, pages 265–278. Springer, 2014.

[16] C. Stauffer and T. Kinh. Automated Multi-Camera Planar Tracking Correspondence Modeling. In *CVPR*, pages I–259–I–266, 2003.

[17] T. Svoboda, D. Martinec, and T. Pajdla. A Convenient Multicamera Self-calibration for Virtual Environments. *Presence: Teleoper. Virtual Environ.*, 14(4):407–422, 2005.

[18] Z. Taylor and J. Nieto. ”Motion-Based Calibration of Multimodal Sensor Arrays. In *ICRA*, pages 4843–4850, 2015.

[19] Z. Taylor and J. Nieto. Motion-Based Calibration of Multimodal Sensor Extrinsics and Timing Offset Estimation. *IEEE Transactions on Robotics*, 32(5):1215–1229, 2016.

[20] C. Zhang and Z. Zhang. Calibration Between Depth and Color Sensors for Commodity Depth Cameras. In *ICME, 2011*, pages 1–6, 2011.

[21] Z. Zhang. A Flexible New Technique for Camera Calibration. *Pattern Analysis and Machine Intelligence*, 22(11):1330–1334, 2000.